

To: "Jordan Hoffner" <jhoffner@google.com>, "Julia Peker" <juliap@google.com>, "Palash Nandy" <palash@youtube.com>
From: "Shashi Seth" <shashis@google.com>
Cc:
Bcc:
Received Date: 2007-05-17 17:49:28 GMT
Subject: First peek at Numbers

Jordan:

Wanted to start sharing some of the early numbers we have got back. Julia has put in an incredible effort and pulled the early data for us - we have not started adding any other signals to the mix yet - this is based on pure view-counts.

Shashi

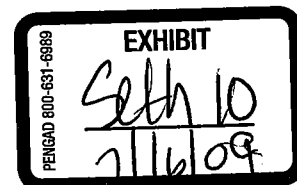
----- Forwarded message -----
From: Julia Peker <juliap@google.com>
Date: May 16, 2007 5:46 PM
Subject: Re: First peek at Numbers
To: shashis@google.com

Hey Shashi,
so i added some numbers to the spreadsheet (check it out). I'm still running % for 10k users (i'm also realizing that my 10K have outliers - people who're gaming the system. There are a couple of people who have over 300K plays in one week - should we automatically exclude them from this analysis?)
i should have the rest of the numbers for you tomorrow ...

thnx,
julia

On 5/16/07, Julia Peker <juliap@google.com > wrote:

>
> Hey Shashi,
> attached are some numbers. These are the grand totals - all views for the
> week of May 6th (April 29th thru May 6th). I have total views, views for
> Premium partners only, and views for content that was taken down for
> copyright issues (some interesting numbers there!)
>
> I'm running 10k users now - will do the similar buckets (i already have
> 10K for total views, running 10K without copyrighted content, and then will
> run both sets with % of UGC vs Premium)
>
> let me know if you have questions!
> j
>
> On 5/16/07, Shashi Seth <shashis@google.com> wrote:
>>
>> Thats perfectly fine to just go with what we have. We can make it
>> perfect in future iterations.
>>
>> On 5/16/07, Julia Peker <juliap@google.com> wrote:
>>>
>>> heh - not as sorry as me - spent way too much time cleaning the data



>>> we were given ;(and now am realizing they definitely didn't give us all
>>> accounts (i know for a fact we have 3 BBC accounts, and they only gave us
>>> 2). But, this is the data we're going with - i really can't spend time on
>>> tracking in down right now ...
>>>
>>> On 5/16/07, Shashi Seth <shashis@google.com> wrote:
>>>>
>>>> Sorry to hear that. Really looking forward to the numbers.
>>>>
>>>> In the meanwhile, is it possible to get our users ranked and get a
>>>> sample of the top 10,000 of them based on their viewcounts? That way once we
>>>> have the content tagged as UGC vs. Premium, then we can also run the report
>>>> from the user side seeing what % of their views are premium vs. UGC.
>>>>
>>>> Thanks
>>>>
>>>> Shashi
>>>>
>>>> On 5/15/07, Julia Peker <juliap@google.com> wrote:
>>>>>
>>>>> well, as my idea of using someone else's work (mainly, someone
>>>>> else telling me who premium partners are) didn't work, i'm working on
>>>>> creating the table myself, and will run the number as soon as i have the
>>>>> table. Just wanted to make sure you know i'm working on it :)
>>>>>
>>>>> j
>>>>>
>>>>> On 5/14/07, Julia Peker <juliap@google.com> wrote:
>>>>>>
>>>>>> cool - thanks, Shashi!
>>>>>> i just emailed with her - she's on it. I'll keep you posted.
>>>>>>
>>>>>> On 5/14/07, Shashi Seth <shashis@google.com> wrote:
>>>>>>>
>>>>>>> Hey:
>>>>>>>
>>>>>>> We are not including Torso partners in this study - yet. Let
>>>>>>> me know if you want me to speak to Shuehan directly to speed things up.
>>>>>>>
>>>>>>> Thanks
>>>>>>>
>>>>>>> Shashi
>>>>>>>
>>>>>>> On 5/14/07, Julia Peker <juliap@google.com> wrote:
>>>>>>>>
>>>>>>>> Hey Shashi,
>>>>>>>> I'm still waiting for Shauhan (sp?) to create the table that
>>>>>>>> holds premium partner info. Remind me if we decided that torso partners
>>>>>>>> should be included in 'premium' partner category for this analysis?
>>>>>>>>
>>>>>>>> i'll follow up with Shauhan to see where she is in creating
>>>>>>>> the table ...
>>>>>>>>
>>>>>>>> thanks,
>>>>>>>> julia
>>>>>>>>
>>>>>>>> On 5/13/07, Shashi Seth <shashis@google.com> wrote:
>>>>>>>>>

```

>>>>>>>> Hi:
>>>>>>>>
>>>>>>>> Just wondering if you will have a chance to pull the first
>>>>>>>> set of numbers by Monday or Tuesday. The first set of numbers are:
>>>>>>>>
>>>>>>>> - For a given week, find all premium content and compute
>>>>>>>> their total views, find all UGC content and compute their total views
>>>>>>>> (remove the premium content from this mix and add it back to premium
>>>>>>>> content)
>>>>>>>>
>>>>>>>> - For the user side, find top 10,000 users based on their
>>>>>>>> video views, and see what content they have watched (premium vs. UGC - using
>>>>>>>> same methodology as above).
>>>>>>>>
>>>>>>>> Please let me know if I can help in any way.
>>>>>>>>
>>>>>>>> Thanks
>>>>>>>>
>>>>>>>> Shashi
>>>>>>>>
>>>>>>>>
>>>>>>>>
>>>>>>>>
>>>>>>>> --
>>>>>>>> *****
>>>>>>>> Julia Peker
>>>>>>>> Datawarehouse Analyst
>>>>>>>> YouTube, Google Inc
>>>>>>>>
>>>>>>>>
>>>>>>>>
>>>>>>>>
>>>>>>>> --
>>>>>>>> *****
>>>>>>>> Julia Peker
>>>>>>>> Datawarehouse Analyst
>>>>>>>> YouTube, Google Inc
>>>>>>>>
>>>>>>>>
>>>>>>>>
>>>>>>>>
>>>>>>>> --
>>>>>>>> *****
>>>>>>>> Julia Peker
>>>>>>>> Datawarehouse Analyst
>>>>>>>> YouTube, Google Inc
>>>>>>>>
>>>>>>>>
>>>>>>>>
>>>>>>>>
>>>>>>>> --
>>>>>>>> *****
>>>>>>>> Julia Peker
>>>>>>>> Datawarehouse Analyst
>>>>>>>> YouTube, Google Inc
>>>>>>>>
>>>>>>>>

```

> >
> >
>
>
> --
> *****
> Julia Peker
> Datawarehouse Analyst
> YouTube, Google Inc
>
>

--

Julia Peker
Datawarehouse Analyst
YouTube, Google Inc

Attachments:

premium_vs_ugc.xls

	Views	Count Users	Avg Views/User	Count Videos	Avg Views/Video	% to Total Views
All	2311629555	1794395468	1.288249773	20932506	110.432528	1
Premium	18524526	14562981	1.27202844	13815	1340.899457	0.008013622
Copyrighted	23678945	18030947	1.313239122	30934	765.4666387	0.010243399